



On the linear independence of shifted powers

Pascal Koiran, Timothée Pecatte, Ignacio García-Marco

► To cite this version:

Pascal Koiran, Timothée Pecatte, Ignacio García-Marco. On the linear independence of shifted powers. Journal of Complexity, 2018, 45, pp.67-82. hal-01520724v2

HAL Id: hal-01520724

<https://hal.science/hal-01520724v2>

Submitted on 20 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the linear independence of shifted powers

Pascal Koiran, Timothée Pecatte

LIP*, Ecole Normale Supérieure de Lyon, Université de Lyon

Ignacio García-Marco[†]

I2M, Aix-Marseille Université

October 18, 2017

Abstract

We call *shifted power* a polynomial of the form $(x - a)^e$. The main goal of this paper is to obtain broadly applicable criteria ensuring that the elements of a finite family F of shifted powers are linearly independent or, failing that, to give a lower bound on the dimension of the space of polynomials spanned by F . In particular, we give simple criteria ensuring that the dimension of the span of F is at least $c \cdot |F|$ for some absolute constant $c < 1$. We also propose conjectures implying the linear independence of the elements of F . These conjectures are known to be true for the field of real numbers, but not for the field of complex numbers. The verification of these conjectures for complex polynomials directly imply new lower bounds in algebraic complexity.

Keywords: linear independence; wronskian; shifted differential equation; real and complex polynomial; Birkhoff interpolation; Waring rank

MSC2010: 12D99, 26A75, 30A06

1 Introduction

In this article, we consider families of univariate polynomials of the form:

$$F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\},$$

where $e_i \in \mathbb{N}$ and the a_i belong to a field \mathbb{K} of characteristic 0. An element of F will be called a *shifted power* (polynomials of this form are also called *affine*

*UMR 5668 Ecole Normale Supérieure de Lyon, CNRS, UCBL, INRIA. Emails: [Pascal.Koiran, Timothee.Pecatte]@ens-lyon.fr, iggarcia@ull.es

[†]Corresponding author. Address: Aix-Marseille Université. I2M UMR 7373, Case Courrier 907. 163 avenue de Luminy. 13288 Marseille Cedex 9, France.

powers in [7]). We always assume that $(a_i, e_i) \neq (a_j, e_j)$ for $i \neq j$, i.e. that F does not contain the same element twice.

The main goal of this paper is to obtain broadly applicable criteria ensuring that the elements of F are linearly independent or, failing that, to give a lower bound on the dimension of the space of polynomials spanned by F . For instance, we have the following well-known result for the case of equal exponents.

Proposition 1.1 (Folklore). *For any integer d , for any distinct $(a_i) \in \mathbb{K}^{d+1}$, the family $\{(x - a_0)^d, \dots, (x - a_d)^d\}$ is a basis of the space of polynomials of degree at most d .*

This can be shown for instance by checking that the Wronskian determinant of the family is not identically 0. Nullity of the Wronskian is a necessary and sufficient condition for the linear independence of polynomials [3, 4, 12], so our problem always reduces in principle to the verification that the Wronskian of F is nonzero. Unfortunately, the resulting determinant looks hardly manageable in general. As a result, little seems to be known in the case of unequal exponents (the case of equal exponents is tractable because the Wronskian determinant becomes a Vandermonde matrix after multiplication of rows by constants). One exception is the so-called Jordan's lemma [8] (see [9, Lemma 1.35] for a recent reference), which provides the following generalization of Proposition 1.1:

Lemma 1.2. *Let $d \in \mathbb{Z}^+$, $e_1, \dots, e_t \in \{1, \dots, d\}$, and let $a_1, \dots, a_t \in \mathbb{K}$ be distinct constants. If $\sum_{i=1}^t (d + 1 - e_i) \leq d + 1$, then the elements of*

$$\bigcup_{i=1}^t \left\{ (x - a_i)^{e_i}, (x - a_i)^{e_i+1}, \dots, (x - a_i)^d \right\}$$

are linearly independent.

So far, we have only discussed sufficient conditions for linear independence. The following “Pólya condition” is an obvious *necessary condition*:

Definition 1.3. *For a sequence $e = (e_1, \dots, e_s)$ of integers, let $n_i = |\{j : e_j < i\}|$. We say that e satisfies the Pólya condition if $n_i \leq i$ for all i .*

For a family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$, we say that F satisfies the Pólya condition if $e = (e_1, \dots, e_s)$ does.

The name *Pólya condition* is borrowed from the theory of Birkhoff interpolation [15, 6]. This necessary condition for linear independence is not sufficient: for instance we have the linear dependence relation $(x + 1)^2 - (x - 1)^2 - 4x = 0$. As we shall see later, the Pólya condition turns out to be sufficient in a probabilistic sense: if the shifts a_i are taken uniformly at random, the resulting family is

linearly independent with high probability. As pointed out above, little is known about deterministic sufficient conditions for linear independence. But there is an exception when \mathbb{K} is the field of real numbers: in this case, some recent progress was made in [6] thanks to a connection between Birkhoff interpolation and linear independence of shifted powers.

In particular, the authors showed that the Pólya condition is only a factor of 2 away from being also a sufficient condition for linear independence:

Theorem 1.4 (Theorem 3 in [6]). *Let F and the n_i 's be as in Definition 1.3, and let $d = \max e_i$. If all the a_i 's are real, and $n_1 \leq 1, n_j + n_{j+1} \leq j + 1$ for all $j = 1 \dots d$, then the elements of F are linearly independent.*

They also gave an example of linear dependence that violates only one of the inequalities of Theorem 1.4, showing that this result is essentially optimal.

Theorem 1.4 fails badly over the field of complex numbers, as shown by this example from [6].

Proposition 1.5. *Take $k \in \mathbb{Z}^+$ and let ξ be a k -th primitive root of unity. Then, for all $d \in \mathbb{Z}^+$ and all $\mu \in \mathbb{C}$ the following equality holds:*

$$\sum_{j=1}^k \xi^j (x + \xi^j \mu)^d = \sum_{\substack{i \equiv -1 \pmod{k} \\ 0 \leq i \leq d}} k \binom{d}{i} \mu^i x^{d-i}. \quad (1)$$

Indeed, if we take $k = \sqrt{d}$ in (1), Proposition 1.5 shows that the $2k = 2\sqrt{d}$ shifted powers in the set

$$\{(x + \xi^j \mu)^d \mid 1 \leq j \leq k\} \cup \{x^{d-i} \mid i \equiv -1 \pmod{k}, 0 \leq i \leq d\}$$

are linearly dependent; and the exponents of these shifted powers clearly satisfy the hypothesis of Theorem 1.4. The question of finding a “good” sufficient condition for linear independence over \mathbb{C} was left open in [6]. We propose the following conjectures.

Conjecture 1.6. *There are absolute constants a and b such that for all large enough integers s , the elements in any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ of s complex shifted powers are linearly independent if $e_i \geq as + b$ for all i .*

If this conjecture holds true, one must have $a \geq 1$. Indeed, a family where $e_i \leq s - 2$ for all i will violate the Pólya condition. One can say more. Indeed, we will see in Proposition 3.1 that the counterpart of Conjecture 1.6 for the field of real numbers holds true, and that one may take $a = 2, b = -4$. We will also show that this result is best possible over \mathbb{R} : one cannot take $a = 2, b = -5$. Therefore, if Conjecture 1.6 holds true one must in fact have $a \geq 2$.

Conjecture 1.7. *There are absolute constants a and b such that for any linearly independent family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ with $s \leq ad + b$, $a_i \in \mathbb{C}$ and $e_i \leq d$ for all i , the elements in the family $F \cup \{(x + 1)^{d+1}, x^{d+1}\}$ are linearly independent as well.*

By Theorem 13 from [6], the counterpart of this conjecture for the field of real numbers holds true. We will not prove any of these two conjectures in the present paper, but we will show that a weak version of Conjecture 1.6 holds true: it suffices to replace the hypothesis $e_i \geq as + b$ by the stronger assumption $e_i \geq s(s - 1)/2$. Moreover, we show that any family of s shifted powers satisfying the Pólya condition spans a space of dimension at least $s/4$ (this should be compared to the straightforward \sqrt{s} lower bound).

Complexity-theoretic motivation

Studying the linear independence of shifted powers is a natural and challenging problem in its own right, but it also has a complexity-theoretic motivation. The motivating problem is to prove lower bounds on the size of the representation of “explicit” polynomials as linear combinations of shifted powers. More precisely, we want a lower bound on the number s of shifted powers needed to represent a polynomial f of degree d as:

$$f = \sum_{i=1}^s \alpha_i (x - a_i)^{e_i}. \quad (2)$$

Clearly, one cannot hope for more than a lower bound that is linear in d , and this was achieved in [6] for real polynomials. For complex polynomials, $\Omega(\sqrt{d})$ lower bounds (i.e., lower bounds of the form $c\sqrt{d}$ for some absolute constant $c > 0$) were obtained for the more general model of sums of powers of low degree polynomials in [10].

The connection between this complexity-theoretic problem and the linear independence of shifted powers arises when the “hard polynomial” f itself is defined as a linear combination of shifted powers. In this case, (2) can be rewritten as a linear dependence relation between shifted powers (those on the right-hand side of (2), and those occurring in the definition of f). If we can show that such a linear dependence is impossible for small enough s , we have a lower bound on s . This is exactly the way that the lower bounds in [6] were obtained. Likewise, Conjecture 1.7 is equivalent to the following complexity-theoretic conjecture at the end of [6]: the polynomial $f(x) = (x + 1)^{d+1} - x^{d+1}$ requires $s = \Omega(d)$ terms to be represented under form (2) with all exponents $e_i \leq d$. The same reasoning shows

that a proof of Conjecture 1.6 would yield an $\Omega(d)$ lower bound for representations of degree d polynomials satisfying constraints of the form $e_i = \Omega(d)$ for all i .

Finally, we note that lower bounds on the dimension of the space spanned by a family F of s shifted powers do not seem to imply any complexity-theoretic lower bound if they fall short of full linear independence ($\dim F = s$). They are hopefully a step in this direction, however.

Organization of the paper

In Section 2 we introduce one of the main tools of this paper: shifted differential equations (or SDE for short). Those are linear differential equations with polynomial coefficients. The name SDE is a reference to Neeraj Kayal’s lower bound method of shifted partial derivatives [16]. We have already used shifted differential equations in the design of efficient algorithms for shifted powers [7]. The main novelty in the present paper is a technical lemma (see Proposition 2.2 and Corollary 2.5) which helps to deal with families where some nodes a_i are repeated, i.e., occur in several shifted powers of the family. This lemma is then used for the study of linear independence (in Section 3) and dimension lower bounds (in Section 4) over the field of complex numbers.

Section 3 provides sufficient conditions for linear independence. We first show that a version of Conjecture 1.6 with $e_i \geq 2s - 4$ holds over the field of real numbers, and that this bound is optimal. For the field of complex numbers, we prove a version of this conjecture under the stronger assumption $e_i \geq s(s - 1)/2$. Finally we show that when the shifts a_i are chosen at random, then with high probability the corresponding family of shifted powers is linearly independent for *all* choices of exponents e_i satisfying the Pólya condition. As an intermediate step, we give an exact formula for the number of Pólya sequences where the s exponents e_i are all bounded by the same integer d .

In Section 4 we give lower bounds on the dimension of the space of polynomials spanned by a family of s shifted powers. We first present a straightforward \sqrt{s} lower bound, and then give lower bounds that are linear in s for the fields of real and complex numbers. We note that the complex lower bound is nonconstructive, i.e., the proof does not yield an explicit independent subfamily of dimension $\Omega(s)$.

2 Shifted Differential Equations

2.1 Definitions

Shifted differential equations are a particular type of differential equations with polynomial coefficients.

Definition 2.1. A Shifted Differential Equation (SDE) of parameters (t, k, l) is a differential equation of the form

$$\sum_{i=0}^k P_i(x) f^{(i)}(x) = 0$$

where f is the unknown function and the P_i are polynomials in $\mathbb{K}[x]$, with $\deg(P_i) \leq i + l$ for all $i \in \{0, \dots, t-1\}$, $\deg(P_i) \leq l$ for all $i \in \{t, \dots, k\}$ and $P_k \neq 0$. We refer to the polynomials P_i as the coefficients of the SDE.

The quantity k is called the order of the equation and the quantity l is called the shift. We will usually denote such a differential equation by $\text{SDE}(t, k, l)$.

This class of differential equations includes the well studied Fuchsian differential equations. A Fuchsian equation is a linear differential equation in which the singular points of the coefficients are regular singular points; in particular, these equations are linear differential equations with polynomial coefficients of bounded degree (see, e.g., [17]).

We want to highlight that the definition of $\text{SDE}(t, k, l)$ also covers the case when $t > k$; thus the SDE defined here are a generalization of those used in [10].

As an example, the shifted power $(x - a)^e$ satisfies the SDE:

$$(x - a)f' - e.f = 0,$$

of order 1 and shift 1 with either $t = 0$ or $t = 1$ (this differential equation can also be seen as a SDE of order 1 and shift 0 with $t = 2$). We generalize this observation to several shifted powers in Proposition 2.2 below.

Shifted differential equations have already been used in [10] to obtain lower bounds and in [7] to obtain efficient algorithms. In these two papers the authors worked with a slightly less general definition of SDEs: they considered SDEs of order k , shift l and the degrees of the P_i 's were all of the form $\deg P_i \leq i + l$. This corresponds now to the choice of parameters $(k + 1, k, l)$ in Definition 2.1, i.e., with $t = k + 1$. The main advantage of adding the new parameter t is that if we have a $\text{SDE}(t, k, l)$ which is satisfied by t linearly independent polynomials, we can ensure that the polynomial P_k has low degree; indeed, its degree is at most l (this should be compared with the value $k + l$ of the previous model of SDEs). The fact that P_k has low degree will play an essential role in some proofs of this work.

The method of shifted differential equations is quite robust, and this allows us to obtain a key ingredient similar to that of [7]: for a family of s shifted powers, there exists a “small” SDE satisfied by all the terms. More precisely:

Proposition 2.2. *Let $F = \{(x - a_i)^{e_i} : a_i \in \mathbb{K}, e_i \in \mathbb{N}, 1 \leq i \leq s\}$. Then for any choice of parameters (t, k, l) such that*

$$s(l + k + 1) < (k + 1)(l + 1) + \frac{t(t - 1)}{2}$$

there exists a SDE(t, k', l) with $k' \leq k$ satisfied simultaneously by the $f_i(x) = (x - a_i)^{e_i}$ for $i = 1, \dots, s$.

Proof. The existence of this common SDE is equivalent to the existence of a common nonzero solution for the following equations $(E_r)_{1 \leq r \leq s}$ in the unknowns $\lambda_{i,j}$:

$$\sum_{i=0}^{t-1} \left(\sum_{j=0}^{i+l} \lambda_{i,j} x^j \right) f_r^{(i)}(x) + \sum_{i=t}^k \left(\sum_{j=0}^l \lambda_{i,j} x^j \right) f_r^{(i)}(x) = 0 \quad (E_r)$$

Therefore, there are $(k + 1)(l + 1) + \frac{t(t-1)}{2}$ unknowns, so we need to show that the matrix of this linear system has rank smaller than $(k + 1)(l + 1) + \frac{t(t-1)}{2}$. We are going to show that for each fixed value of $r \in \{1, \dots, s\}$, the subsystem (E_r) has a matrix of rank $\leq l + k + 1$. In other words, we have to show that the subspace V_r has dimension less than $l + k + 1$, where V_r is the linear space spanned by the polynomials $x^j f_r^{(i)}(x)$, with $(i, j) \in \mathbb{N}^2$ such that either $0 \leq i \leq t - 1$ and $0 \leq j \leq i + l$; or $t \leq i \leq k$ and $0 \leq j \leq l$. But V_r is included in the subspace spanned by the polynomials

$$\{(x - a_r)^{e_r+j}; -k \leq j \leq l, e_r + j \geq 0\}.$$

This is due to the fact that the polynomials x^i belongs to the span of the polynomials $\{(x - a_r)^\ell : 0 \leq \ell \leq i\}$. Hence, we have that $\dim V_r \leq l + k + 1$. Since the (E_r) subsystem has a matrix of rank $\leq l + k + 1$, the whole system has rank $\leq s(l + k + 1)$. Thus, there exists a nonzero solution if $(k + 1)(l + 1) + \frac{t(t-1)}{2} > s(l + k + 1)$. \square

Notice that if we assume that the elements of F are linearly independent, it cannot be part of the solution set of a differential equation of order strictly less than s . Thus the differential equation we obtain from previous Proposition is a SDE(t, k', l) with $s \leq k' \leq k$. In particular $\deg(P_{k'}) \leq l$ if we choose $t \leq s$, therefore we will usually set $t = s$ in the following.

In the remainder of Section 2.1, we discuss a few choices of interest of the pair (k, l) for a given set of shifted powers F of size s . Notice first that the expression $s(l + k + 1) < (k + 1)(l + 1) + t(t - 1)/2$ in Proposition 2.2 is symmetric in k and l .

Case $l < s$: we obtain that there exists a $\text{SDE}(s, k, s-1)$ if $s^2 < s$, which is not possible. The fact that in general one cannot find a $\text{SDE}(s, k, l)$ with $l < s$ satisfied by a family of s affine powers will be justified in Remark 2.6.

Case $k = s$: we obtain $l \geq s(s+1)/2$. This is similar to the SDE coming from the factorized Wronskian exhibited in the first version of [7].

Case $l = s$: we obtain $k \geq s(s+1)/2$. This choice of parameters yields a result of linear independence over \mathbb{C} , as we will prove in Proposition 3.5.

Case $k = \alpha \cdot s$ with $\alpha > 1$: it is enough to have $l \geq \frac{2\alpha-1}{2\alpha-2} s$, which shows the tradeoff between k and l . The “point in the middle” where $k = l$ corresponds to $k = l = \left(1 + \frac{\sqrt{2}}{2}\right) s$.

2.2 Roots of coefficients of a differential equation

We now proceed to one of the main tools of this paper. We show that a differential equation with polynomial coefficients satisfied by every element of a family F of shifted powers must have some structure in the roots of its coefficients. In this section we will use the convenient notation:

$$x^{\underline{i}} = x(x-1) \cdots (x-i+1)$$

where i is a positive integer (for $i = 0$ we set $x^{\underline{i}} = 1$). When x is a nonnegative integer, we have $x^{\underline{i}} = x!/(x-i)!$ for $x \geq i$ and $x^{\underline{i}} = 0$ for $x < i$. This notation allows us to write the i -th derivative of a shifted power $f(x) = (x-a)^e$ in a concise way: $f^{(i)} = e^{\underline{i}}(x-a)^{e-i}$.

We first make the following remark, which was already present in [7].

Proposition 2.3. *Consider the following differential equation with polynomial coefficients:*

$$\sum_{i=0}^k P_i(x) f^{(i)}(x) = 0.$$

Assume that $(x-a)^e$ satisfies this equation, with $e \geq k$. Then we have $P_k(a) = 0$.

Proof. Since $(x-a)^e$ is a solution, we have

$$\sum_{i=0}^k P_i(x) e^{\underline{i}} (x-a)^{e-i} = 0.$$

We deduce that there exists $q \in \mathbb{K}[x]$ such that $P_k(x)(x-a)^{e-k} = (x-a)^{e-k+1}q(x)$, from which we deduce that $P_k(a) = 0$. \square

As a direct consequence, if $(x-a)^e$ and $(x-b)^f$ with $a \neq b$ and $e, f \geq k$ both satisfy the same equation, then both a and b are roots of P_k . However, if $(x-a)^e$ and $(x-a)^f$ both satisfy the same equation, can we say more than just $P_k(a) = 0$? We will answer positively this question thanks to the following proposition:

Proposition 2.4. *Let $(*)$ be the following differential equation with polynomial coefficients:*

$$\sum_{i=0}^k P_i(x) f^{(i)}(x) = 0 \quad (*)$$

If $()$ is satisfied simultaneously by x^{e_1}, \dots, x^{e_n} where $n \leq k$ and $e_1 > e_2 > \dots > e_n \geq k - n + 1$, then for all $m = 0 \dots n - 1$, x^{n-m} divides P_{k-m} .*

Proof. By injecting x^{e_j} into $(*)$, we get the following equation:

$$\sum_{i=0}^{\min(e_j, k)} P_i(x) e_j^i x^{e_j-i} = 0$$

If $e_j \leq k$, we multiply the equation by x^{k-e_j} , and otherwise, we factor out x^{e_j-k} . In both cases, we obtain:

$$\sum_{i=0}^k e_j^i \cdot P_i(x) x^{k-i} = 0 \quad (E_j)$$

We have n such equations $(E_j)_{1 \leq j \leq n}$ and we will now take a "good" linear combination to deduce the result.

Fix an integer $0 \leq m < n$, and consider the n -tuple $\vec{u}_m \in \mathbb{K}^n$ such that $(\vec{u}_m)_i = 0$ for all $i \neq n - m$ and $(\vec{u}_m)_{n-m} = 1$. A "good" linear combination is a n -tuple $(\alpha_1, \dots, \alpha_n)$ such that there exists (b_0, \dots, b_{k-n}) satisfying

$$\left(\begin{array}{ccc} e_1^0 & \dots & e_n^0 \\ e_1^1 & \dots & e_n^1 \\ \vdots & \ddots & \vdots \\ e_1^{k-1} & \dots & e_n^{k-1} \\ e_1^k & \dots & e_n^k \end{array} \right) \cdot \left(\begin{array}{c} \alpha_1 \\ \vdots \\ \alpha_n \end{array} \right) = \left(\begin{array}{c} b_0 \\ \vdots \\ b_{k-n} \\ \vec{u}_m \end{array} \right) \left. \vphantom{\begin{array}{c} b_0 \\ \vdots \\ b_{k-n} \\ \vec{u}_m \end{array}} \right\} \begin{array}{l} k - n + 1 \text{ rows} \\ n \text{ rows} \end{array}$$

We claim that it is always possible to find such a tuple. Assuming this fact, we then compute the following equation:

$$0 = \sum_{j=1}^n \alpha_j(E_j) = \sum_{i=0}^{k-n} b_i P_i(x) x^{k-i} + P_{k-m}(x) x^m$$

This directly implies that x^{n-m} divides $P_{k-m}(x)$.

To prove the claim, we will use the proof technique of Lemma 2 from [13] to show that the following square submatrix is invertible:

$$\begin{pmatrix} e_1^{k-n+1} & \dots & e_n^{k-n+1} \\ \vdots & \ddots & \vdots \\ e_1^k & \dots & e_n^k \end{pmatrix}.$$

We first factorize its determinant using the fact that $a^{b+c} = a^b \cdot (a-b)^c$ to obtain

$$\begin{vmatrix} e_1^{k-n+1} & \dots & e_n^{k-n+1} \\ \vdots & \ddots & \vdots \\ e_1^k & \dots & e_n^k \end{vmatrix} = \prod_{i=1}^n e_i^{k-n+1} \cdot \begin{vmatrix} d_1^0 & \dots & d_n^0 \\ \vdots & \ddots & \vdots \\ d_1^{n-1} & \dots & d_n^{n-1} \end{vmatrix}$$

where $d_i = e_i - k + n - 1$. Notice that the constant we have factorized is non-zero, since $e_i \geq k - n + 1$. Assume for contradiction that the rows are linearly dependent. This implies that there exists a nonzero tuple (α_i) such that for all $j = 1 \dots n$, we have $\sum_{i=0}^{n-1} \alpha_i d_j^i = 0$. In other words, if we consider the polynomial $P(x) = \sum_{i=0}^{n-1} \alpha_i x^i$, then $P(d_j) = 0$ for all j . However, all the d_j 's are distinct and P is of degree at most $n - 1$, a contradiction. \square

As a consequence, we obtain the following refinement of Proposition 2.3.

Corollary 2.5. *Let $(*)$ be the following differential equation with polynomial coefficients of order $k > 0$:*

$$\sum_{i=0}^k P_i(x) f^{(i)}(x) = 0 \tag{*}$$

Consider a family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s, e_i \geq k\}$ such that $()$ is satisfied simultaneously by all the elements of F . Then $\prod_{i=1}^s (x - a_i)$ divides P_k .*

In other words, each node a_i is a root of P_k of multiplicity greater or equal to the number of occurrences of this node in the family F .

Proof. We partition F in subfamilies along the values of the a_i 's: $F = \uplus_{i=1}^t F_i$ where $F_i = \{(x - b_i)^{\varepsilon_{i,j}} : 1 \leq j \leq s_i\}$, such that $b_i \neq b_j$ for $i \neq j$. Notice

that $\prod_{i=1}^s (x - a_i) = \prod_{i=1}^t (x - b_i)^{s_i}$, and since $b_i \neq b_j$, it is enough to show that for $i = 1 \dots t$, we have $(x - b_i)^{s_i}$ divides P_k . This is obtained directly by using Proposition 2.4 with $m = 0$. \square

Remark 2.6. *This explains why in general one cannot find a SDE satisfied simultaneously by a family of s shifted powers with shift $l < s$. Indeed, for any choice of parameters (s, k, l) with $k \geq s$, consider a family of shifted powers $F = \{(x - a_i)^k : 1 \leq i \leq s\}$. Take any SDE (s, k', l) with $k' \leq k$ satisfied by F . Since the elements of F are linearly independent, then $k' \geq s$. Hence, Corollary 2.5 yields that the polynomial $\prod_{i=1}^s (x - a_i)$ divides $P_{k'}$, which has degree at most l , implying $l \geq s$.*

We now remove the hypothesis of “big exponents” ($e_i \geq k$) and prove that every node in the family should appear as a root of one of the coefficients of the differential equation.

Corollary 2.7. *Let $(*)$ be the following differential equation with polynomial coefficients:*

$$\sum_{i=0}^k P_i(x) f^{(i)}(x) = 0. \quad (*)$$

with $P_0 \neq 0$. Define $I = \{i : P_i(x) \neq 0\}$. Consider a family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ such that $()$ is satisfied simultaneously by all the elements of F . Then $\prod_{i=1}^s (x - a_i)$ divides $\prod_{i \in I} P_i$.
Indeed, for a given index i , $(x - a_i)$ will divide P_j with $j = \max\{p : p \in I, p \leq e_i\}$.*

Proof. Without loss of generality, we can assume that $e_1 \geq e_2 \geq \dots \geq e_s$, and we write $F = \{f_1, \dots, f_s\}$ where $f_i = (x - a_i)^{e_i}$. We consider the last index p such that $e_p \geq k$ and partition F into two sets: $F = \{f_1, \dots, f_p\} \cup \{f_{p+1}, \dots, f_s\}$. Using Corollary 2.5, we get that $\prod_{i=1}^p (x - a_i)$ divides P_k . We now consider the following equation:

$$\sum_{i=0}^{k-1} P_i(x) f^{(i)}(x) = 0 \quad (*')$$

Notice that for $i > p$, f_i satisfies $(*)'$. Since the order of the equation has decreased, and since $0 \in I$, we can proceed by induction to obtain that $\prod_{i=p+1}^s (x - a_i)$ divides $\prod_{i \in I \setminus \{k\}} P_i$. The combination of these two facts yields the wanted result. \square

3 Linear independence

The goal of this section is to prove sufficient conditions for the linear independence of a family of shifted powers. One trivial such condition is when the degrees of the polynomials in the family are all distinct. As pointed out in Proposition 1.1, linear independence also holds in the case of equal exponents. In this section we provide further sufficient conditions for linear independence and study their optimality.

3.1 The real case

In the following we derive from Theorem 1.4 another sufficient condition for the linear independence of families without small exponents, and show that the result is tight.

Proposition 3.1. *For any family $F = \{(x - a_i)^{e_i} : a_i \in \mathbb{R}, e_i \geq \max(1, 2s - 4), 1 \leq i \leq s\}$, the elements of F are linearly independent.*

Proof. Assume without loss of generality that $d = e_1 \geq e_2 \geq \dots \geq e_s$. We first eliminate a few trivial cases:

- $s = 1$: the elements of F are linearly independent.
- $s = 2$: two affine powers are linearly dependent if and only if they are equal, thus the elements of F are linearly independent.
- $e_1 = e_2 = \dots = e_s$. Then all the a_i are distinct because all (e_i, a_i) are distinct. Since $d \geq 2s - 4$ and $s \geq 3$, we have $s \leq d + 1$ and the elements of F are linearly independent by Proposition 1.1.
- $e_1 > e_2$: no linear dependence could involve $(x - a_1)^{e_1}$, hence it is enough to show that the elements of the subfamily $F' = \{(x - a_i)^{e_i} : 2 \leq i \leq s\}$ are linearly independent. Since F' satisfies the hypotheses of the Proposition, we can deal with this case by induction on s .

We now have $e_1 = e_2 > e_s$ and $s \geq 3$ which implies that $d \geq 2s - 3$ and that $n_d \leq s - 2$. Such a family satisfies the hypotheses of Theorem 1.4, which directly yields the result. Indeed:

- For $i = 0$, we have $n_1 = 0 \leq 1$.
- For $i \leq 2s - 5$, we have $n_i + n_{i+1} = 0$.
- For $2s - 5 < i < d$, we have $n_i + n_{i+1} \leq 2(s - 2) \leq i + 1$.
- For $i = d$, we have $n_d + n_{d+1} \leq 2s - 2 \leq d + 1$.

□

In order to show that the bound $e_i \geq 2s - 4$ in the above result is tight, we will consider the real polynomial $H_{2d+1}(x) = (x+1)^{2d+2} - x^{2d+2}$ and write it as a sum of $(2d+1)$ -st powers. Define the Waring rank $\text{Waring}_{\mathbb{K}}(f)$ of a univariate polynomial f as

$$\min \left\{ s : \exists \alpha \in \mathbb{K}^s, a \in \mathbb{K}^s, f = \sum_{i=1}^s \alpha_i (x + a_i)^{\deg f} \right\} \quad (3)$$

Waring rank has been studied by algebraists and geometers since the 19th century. It is classically defined for a homogeneous polynomial of degree d as the smallest number s such that f can be written as a sum of s d -th powers of linear forms. The Waring rank of a univariate polynomial $f \in \mathbb{K}[X]$ as defined in (3) is closely related with the Waring rank of $f_H \in \mathbb{K}[x, y]$, its homogenization with respect to the variable y . Indeed, it is straightforward to check that:

- $\text{Waring}_{\mathbb{K}}(f_H) \leq \text{Waring}_{\mathbb{K}}(f)$, and
- if $f_H = \sum_{i=1}^w (a_i x + b_i y)^d$ with $a_i, b_i \in \mathbb{K}$ and $a_i \neq 0$ for all i , then $\text{Waring}_{\mathbb{K}}(f) \leq w$.

The algorithmic question of computing the Waring rank of bivariate complex homogeneous polynomials was solved by Sylvester in 1852. We refer to [9] for the historical background and in the following proof we will use the algorithm described more recently in [5]. We now use this algorithm to show that H_{2d+1} has a “large” Waring rank:

Proposition 3.2. *We have $\text{Waring}_{\mathbb{R}}(H_{2d+1}) = \text{Waring}_{\mathbb{C}}(H_{2d+1}) = d + 1$.*

Proof. We will use the algorithmic result in [5, Section 3] to compute the complex Waring rank of H_{2d+1} and then prove that it coincides with its real Waring rank. We consider $P(x, y)$ the homogenization of H_{2d+1} with respect to the variable y : $P(x, y) = \sum_{i=0}^{2d+1} \binom{2d+2}{i} x^i y^{2d+1-i} = \sum_{i=0}^{2d+1} \binom{2d+2}{i+1} x^{2d+1-i} y^i$. We extract the coefficients $Z_i = \frac{\text{coeff}(P, x^{2d+1-i} y^i)}{\binom{2d+1}{i}} = \frac{\binom{2d+2}{i+1}}{\binom{2d+1}{i}} = \frac{2d+2}{i+1}$ and following [5], we construct the matrix

$$M = \begin{pmatrix} Z_0 & Z_1 & \cdots & Z_d \\ Z_1 & Z_2 & \cdots & Z_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{d+1} & Z_{d+2} & \cdots & Z_{2d+1} \end{pmatrix} = (2d+2) \cdot \begin{pmatrix} \frac{1}{1} & \frac{1}{2} & \cdots & \frac{1}{d+1} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{d+2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{d+2} & \frac{1}{d+3} & \cdots & \frac{1}{2d+2} \end{pmatrix}$$

The last matrix is a Hilbert matrix with an additional row. Hilbert matrices are known to be invertible, as special cases of Cauchy matrices, therefore we have $\text{rank}(M) = d + 1$, which implies, according to [5] that the complex Waring rank is either $d + 1$ or $d + 2$. In order to show that it is in fact $d + 1$, we have to find a vector $f \in \mathbb{C}^{d+2}$ in the kernel of M^t (which is unique up to scalar multiplication) and prove that the corresponding polynomial $F(x) = \sum_{i=0}^{d+1} f_i x^i$ does not have multiple roots.

Notice that the i^{th} row of $M^t f$ can be rewritten as

$$(M^t f)_i = \sum_{j=0}^{d+1} \frac{1}{i+j+1} f_j = \sum_{j=0}^{d+1} \int_0^1 x^i \cdot f_j x^j dx = \int_0^1 x^i F(x) dx$$

The equality $M^t f = 0$ can thus be restated as $\langle F, x^i \rangle = 0$ for $i = 0 \dots d$, with the corresponding scalar product $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$. Such a polynomial can be obtained by the Gram-Schmidt process to $\{1, x, \dots, x^{d+1}\}$ and is classically known as the shifted Legendre polynomial: it can be obtained from the Legendre polynomial by the affine transformation $x \mapsto 2x - 1$. A classical result (see, e.g., [1]) is that the Legendre polynomial of degree $d + 1$ has $d + 1$ distinct real roots in the interval $(-1, 1)$. Therefore our polynomial F has $d + 1$ distinct real roots in the interval $(0, 1)$. This shows that $\text{Waring}_{\mathbb{C}}(H) = d + 1$.

Moreover, if we denote by (a_i) the roots of F , there exist coefficients (α_i) such that $P(x, y) = \sum_{i=1}^{d+1} \alpha_i (x - a_i y)^{2d+1}$. Since the a_i are real, if we take the real part of this equality, we obtain $P(x, y) = \sum_{i=1}^{d+1} \Re(\alpha_i) (x - a_i y)^{2d+1}$ proving that $\text{Waring}_{\mathbb{R}}(P) = \text{Waring}_{\mathbb{C}}(P) = d + 1$. Since $a_i \neq 0$ for all i , we conclude that $\text{Waring}_{\mathbb{R}}(H_{2d+1}) = \text{Waring}_{\mathbb{C}}(H_{2d+1}) = d + 1$. \square

Remark 3.3. Up to multiplication of rows and columns by constants, the matrix M in the above proof is nothing but the matrix of d -th order partial derivatives of P . This explains why the Waring rank of P is greater or equal to the rank of M . Again, we refer to [5] for a proof that the Waring rank is in fact equal to $\text{rank}(M)$ or $d + 2 - \text{rank}(M)$.

Remark 3.4. A similar proof shows that $\text{Waring}_{\mathbb{R}}(H_{2d}) = d + 1$, proving that in general $\text{Waring}_{\mathbb{R}}(H_d) = \lceil \frac{d+1}{2} \rceil$.

By this result, there exist $\alpha_1, \dots, \alpha_{d+1} \in \mathbb{R}$ and $a_1, \dots, a_{d+1} \in (0, 1)$ such that

$$(x + 1)^{2d+2} - x^{2d+2} = \sum_{i=1}^{d+1} \alpha_i (x - a_i)^{2d+1}$$

This equality is a linear dependence of $d + 3$ terms of degree at least $2d + 1 = 2(d + 3) - 5$, showing the optimality of the bound $e_i \geq 2s - 4$.

3.2 The complex case

In the complex case, we can prove a similar sufficient condition for linear independence:

Proposition 3.5. *For any family $F = \{(x - a_i)^{e_i} : a_i \in \mathbb{C}, e_i \geq (s - 1)s/2, 1 \leq i \leq s\}$, the elements of F are linearly independent.*

Proof. Take $G \subseteq F$ a minimal generating subfamily of F and suppose for contradiction that $|G| = t < s$. Using Proposition 2.2, there exists a $\text{SDE}(t, k, t)$ for some $k \leq (t + 1)t/2$ satisfied simultaneously by every element of G :

$$\sum_{i=0}^k P_i(x) f^{(i)}(x) = 0$$

Moreover, since $\langle G \rangle = \langle F \rangle$, we have that every element of F satisfies this SDE. Using Corollary 2.5, since $e_i \geq (s - 1)s/2 \geq t(t + 1)/2 \geq k$, we thus have that $\prod_{i=1}^s (x - a_i)$ divides P_k . This yields $s \leq \deg P_k \leq t < s - 1$, a contradiction. \square

We do not know if the bound $e_i \geq (s - 1)s/2$ is tight. The best example we know is the one provided in the real case that achieves a linear dependence of s affine powers with $e_i \geq 2s - 5$.

3.3 Genericity and linear independence

Let \mathcal{P}_s denote the set of sequences $e = (e_1, \dots, e_s) \in \mathbb{N}^s$ satisfying the Pólya condition. The goal of this section is to study two different random processes for generating a family of affine powers, and to bound the probability that the elements of the generated family are linearly independent. In Corollary 3.8, we study the case where the sequence $e \in \mathcal{P}_s$ is fixed, and the a_i 's are taken uniformly and independently from a set S . In Theorem 3.13, we study the case where we take the a_i 's uniformly and independently from a set S and we want them to give independent families for any $e \in \mathcal{P}_s$. Notice that Corollary 3.8 does not directly implies Theorem 3.13 as the number of sequences in \mathcal{P}_s is infinite.

We first prove that given a family F of s shifted powers such that $1 \notin \langle F \rangle$, the number of shifted powers in $\langle F \rangle$ of degree d is upper bounded by a linear expression in s . The condition $1 \notin \langle F \rangle$ is satisfied if the elements of the set of derivatives $F' = \{f' : f \in F\}$ are linearly independent.

Proposition 3.6. *Consider a family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ such that $1 \notin \langle F \rangle$. Let $S(d) = \{(x - a)^d \in \langle F \rangle\}$. Then for any $e \in \mathbb{N}$, we have $|S(e)| \leq 2s - 1$.*

Proof. Notice first that for any $e \in \mathbb{N}$, we must have $|S(e)| \leq e$. Otherwise, it would contain a basis of $\mathbb{K}_e[x]$ and we could obtain 1 as a linear combination of elements of F , which contradicts the hypothesis. Therefore we are done if $e \leq 2s - 1$.

Otherwise, Proposition 2.2 ensures that there exists a $\text{SDE}(0, k, l)$ with $k = l = 2s - 1$ of order $k' \leq k$ satisfied by all the elements of F :

$$\sum_{i=0}^{k'} P_i(x) f^{(i)}(x) = 0 \quad (*)$$

Since this equation is satisfied by all the elements of $S(e)$, we use Proposition 2.5 to obtain: $|S(e)| \leq \deg P_{k'} \leq l \leq 2s - 1$. \square

Given a sequence $e \in \mathcal{P}_s$, we can take the a_i 's uniformly and independently in an iterative way and use the previous result to lower bound the probability that the elements of the resulting family of shifted powers are linearly independent at each step. The resulting bound is given in Corollary 3.8. Its proof requires the following technical lemma.

Lemma 3.7. *Let $e = (e_1, \dots, e_{s+1}) \in \mathcal{P}_{s+1}$, then $f = (e_1 - 1, \dots, e_s - 1) \in \mathcal{P}_s$.*

Proof. We set $n_i := |\{j : e_j < i, 1 \leq j \leq s+1\}|$ and $n'_i := |\{j : e_j - 1 < i, 1 \leq j \leq s\}|$. Since e satisfies the Pólya condition we have that $n_i \leq i$ for all i . Moreover, since e is non-increasing, we have that

- $n'_i = 0$ if $i \leq e_{s+1}$,
- $n'_i = n_{i+1} - 1 \leq i$ if $i > e_{s+1}$;

hence, f also satisfies the Pólya condition. \square

Corollary 3.8. *Let $e = (e_1, \dots, e_s) \in \mathcal{P}_s$ and let S be a finite subset of \mathbb{K} . Let a_1, \dots, a_s be selected at random independently and uniformly from S . Then*

$$\Pr \left[\{(x - a_i)^{e_i} : 1 \leq i \leq s\} \text{ is linearly independent} \right] \geq 1 - \frac{s(s-1)}{|S|}$$

Proof. We will prove this by induction on s : for $s = 1$, we always obtain a linearly independent family. We now consider $e = (e_1, \dots, e_{s+1}) \in \mathcal{P}_{s+1}$, and we define the following events on possible outcomes (a_1, \dots, a_{s+1}) :

$$\begin{aligned} A &= \{ \{(x - a_i)^{e_i} : 1 \leq i \leq s+1\} \text{ is linearly independent} \} \\ B &= \{ \{(x - a_i)^{e_i} : 1 \leq i \leq s\} \text{ is linearly independent} \} \\ C &= \{ \{(x - a_i)^{e_i-1} : 1 \leq i \leq s\} \text{ is linearly independent} \} \end{aligned}$$

Notice that $C \subseteq B \subseteq A$. Using Lemma 3.7 and the induction hypothesis we obtain $\Pr(C) \geq 1 - \frac{s(s-1)}{|S|}$. From Proposition 3.6 we have $\Pr(A|C) \geq 1 - \frac{2s}{|S|}$. Since $\Pr(A) = \Pr(C) \cdot \Pr(A|C)$ we obtain the inequality

$$\Pr(A) \geq \left(1 - \frac{s(s-1)}{|S|}\right) \cdot \left(1 - \frac{2s}{|S|}\right) \geq 1 - \frac{s(s+1)}{|S|}.$$

□

In the remainder of this section, our goal is to prove Theorem 3.13 which states the following: if $a_1, \dots, a_s \in \mathbb{K}$ are selected independently and uniformly at random from a big enough finite set S , then with high probability $(x - a_i)^{e_i} : 1 \leq i \leq s$ are linearly independent for all $e = (e_1, \dots, e_s) \in \mathcal{P}_s$. A key ingredient of this proof is to have a bound on $\max(e_i)$ for a sequence $e \in \mathcal{P}_s$ such that the elements of the family might be linearly dependent. This bound is derived from the following structural result proved in [7].

Proposition 3.9. [7, Proposition 3.11] Suppose $\sum_{i=1}^s \alpha_i (x - a_i)^{e_i} = 0$ with $\alpha_i \neq 0$ and $(a_i, e_i) \neq (a_j, e_j)$ for all i, j . Then $\max(e_i) < \frac{s^2}{2} - 1$.

Motivated by this bound, we naturally define the bounded version of \mathcal{P}_s : $\mathcal{P}'_s = \{e \in \mathcal{P}_s : \max(e_i) \leq \frac{s^2}{2} - 2\}$. The next Corollary ensures that if an outcome (a_1, \dots, a_s) yields a linearly independent family for any $e \in \mathcal{P}'_s$, then it also yields a linearly independent family for any $e \in \mathcal{P}_s$.

Lemma 3.10. Let $e = (e_1, \dots, e_s) \in \mathcal{P}_s$ for $s \geq 2$. If we take f_i such that $\min\{e_i, \frac{s^2}{2} - s\} \leq f_i \leq \frac{s^2}{2} - 2$ for all i , then $f = (f_1, \dots, f_s) \in \mathcal{P}'_s$.

Proof. We have $n'_i = n_i$ for $i \leq \frac{s^2}{2} - s + 1$, and $n'_i \leq s \leq i$ for $i \geq \frac{s^2}{2} - s + 2$. □

Corollary 3.11. We define the following events on possible outcomes (a_1, \dots, a_s) :

$$\begin{aligned} A &= \{\bigwedge_{e \in \mathcal{P}_s} \{(x - a_i)^{e_i} : 1 \leq i \leq s\} \text{ is linearly independent}\} \\ B &= \{\bigwedge_{e \in \mathcal{P}'_s} \{(x - a_i)^{e_i} : 1 \leq i \leq s\} \text{ is linearly independent}\} \end{aligned}$$

Then $A = B$.

Proof. We first observe that if $a_1 = \dots = a_s$, then $A = B$ trivially because $\{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ is linearly independent for every e_1, \dots, e_s , so let us assume that they are not all equal.

Since $\mathcal{P}'_s \subseteq \mathcal{P}_s$, we have $A \subseteq B$. Given an outcome $a = (a_1, \dots, a_s) \in B$ and a sequence $e \in \mathcal{P}_s$, we can distinguish two cases. If $e \in \mathcal{P}'_s$, then the s shifted powers $(x - a_1)^{e_1}, \dots, (x - a_s)^{e_s}$ are linearly independent since $a \in B$.

Otherwise, assume there exists α_i such that $\sum_{i=1}^s \alpha_i (x - a_i)^{e_i} = 0$. We denote by $I = \{i : e_i > s^2/2 - 2\}$, and, using Proposition 3.9, we have that $\alpha_i = 0$ for $i \in I$. We therefore rewrite the equality as $\sum_{i=1}^s \alpha_i (x - a_i)^{f_i} = 0$, where $f_i := e_i$ for $i \notin I$; otherwise, f_i is chosen in $\{s^2/2 - s, \dots, s^2/2 - 2\}$ in such a way that there are no two equal (a_i, f_i) (we observe that we can always choose f_i in this interval since there are $s - 1$ possible values to choose from and there are at least two different a_i 's). Using Lemma 3.10, we have that $f \in \mathcal{P}'_s$ and thus $(x - a_i)^{f_i} : 1 \leq i \leq s$ are linearly independent since $a \in B$, proving that all the α_i 's are zero. \square

Now that we have restricted our attention to a finite set \mathcal{P}'_s , we can directly use the union bound on all possible sequences $e \in \mathcal{P}'_s$ to obtain the result using Corollary 3.8 for a fixed sequence. We only need to obtain an upper bound on $|\mathcal{P}'_s|$; this is done in following Proposition. More precisely, we will compute exactly $|\mathcal{P}_{s,d}|$ where $\mathcal{P}_{s,d} = \{e \in \mathcal{P}_s : \max(e_i) < d\}$.

Proposition 3.12. *Let $s \leq d$ be integers. Then*

$$|\mathcal{P}_{s,d}| = \binom{s+d}{s} \frac{d+1-s}{d+1}$$

Proof. Notice first that a sequence $e \in \mathcal{P}_{s,d}$ can be represented by the d -tuple (m_1, \dots, m_d) , where $m_i = |\{j : e_j = i - 1\}|$. Therefore, there is a bijection between $\mathcal{P}_{s,d}$ and the following set:

$$\mathcal{Q}_{s,d} = \left\{ (m_1, \dots, m_d) \in \mathbb{N}^d : \forall j \leq d, \sum_{i=1}^j m_i \leq j \wedge \sum_{i=1}^d m_i = s \right\}$$

For each $m \in \mathcal{Q}_{s,d}$, we associate a lattice path on the Cartesian plane as follows: start the path at $(0, 0)$, and at the i -th step move right 1 unit then go up m_i units. The resulting path ends at position (d, s) , and never goes above the diagonal $y = x$. In fact, there is a bijection between $\mathcal{Q}_{s,d}$ and the monotonic lattice paths starting at position $(0, 0)$, ending at position (d, s) , and not passing above the diagonal $y = x$. These numbers are usually called *ballot numbers*, and have been studied since de Moivre (1711). The analytic expression can be found in [11, p. 451], proving the result. \square

In the Birkhoff interpolation paper [14] a similar proof technique was used to count the number of Pólya matrices.

Theorem 3.13. *Let S be a finite subset of \mathbb{K} . Let a_1, \dots, a_s be selected at random independently and uniformly from S . Then*

$$\Pr \left[\bigwedge_{e \in \mathcal{P}_s} \{(x - a_i)^{e_i} : 1 \leq i \leq s\} \text{ is linearly independent} \right] \geq 1 - \frac{f(s)}{|S|}$$

where $f(s) = \left(s + \frac{s^2}{2} - 1\right)(s-1)(s-2)$.

Proof. Following the notation of Corollary 3.11, we have $\Pr(A) = \Pr(B)$. We compute $\Pr(B)$ using the union bound:

$$\Pr(B) \geq 1 - \sum_{e \in \mathcal{P}'_s} \Pr(\{(x - a_i)^{e_i} : 1 \leq i \leq s\} \text{ is linearly dependent})$$

By Corollary 3.8 we have $\Pr(B) \geq 1 - |\mathcal{P}'_s| \cdot \frac{s(s-1)}{|S|}$. Using Proposition 3.12 with $d = \frac{s^2}{2} - 1$, we obtain

$$\Pr(B) \geq 1 - \left(s + \frac{s^2}{2} - 1\right) \cdot \frac{\frac{s^2}{2} - s}{\frac{s^2}{2}} \cdot \frac{s(s-1)}{|S|}.$$

□

Remark 3.14. One can improve the previous lower bound by noticing that for some sequences $e \in \mathcal{P}'_s$, we have $\Pr(\{(x - a_i)^{e_i}\} \text{ is linearly dependent}) = 0$. This is the case for instance for any sequence e with distinct e_i 's, thus we have

$$\Pr(B) \geq 1 - \left(|\mathcal{P}'_s| - \binom{s^2/2}{s}\right) \cdot \frac{s(s-1)}{|S|}.$$

4 Dimension lower bounds

As sufficient conditions for linear independence are hard to find, we now try to lower bound the dimension of the span of families of affine powers instead. Of course, this can only be easier since linear independence implies full dimension of the corresponding space. We mostly investigate two different cases: when all the exponents are big, as in Sections 3.1 and 3.2, and when small exponents are allowed. In the latter case, we will require that the family F satisfies the *Pólya condition* from the Introduction. Under this condition, we first prove an easy lower bound which holds over any field.

Proposition 4.1. Consider any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ satisfying the *Pólya condition*. Then we have $\dim\langle F \rangle \geq \sqrt{s}$.

Proof. We partition F according to the values of the exponents: $F = \cup_{i=1}^t F_i$ with $F_i = \{(x - a_{i,j})^{d_i} : 1 \leq j \leq t_i\}$, and $d_i \neq d_j$ for $i \neq j$. The fact that F satisfies the *Pólya condition* implies that $d_i \geq t_i - 1$, and therefore that every F_i is an independent family, using Proposition 1.1.

Therefore, if there exists $k \in [1; t]$ such that F_k contains at least \sqrt{s} elements we have directly $\dim\langle F \rangle \geq \dim\langle F_k \rangle = |F_k| \geq \sqrt{s}$. Otherwise, we must have $t \geq \sqrt{s}$. We consider a family G obtained by taking one element in each F_i . Since the d_i 's are distinct, the elements of G are linearly independent. This proves that $\dim\langle F \rangle \geq |G| = t \geq \sqrt{s}$. \square

In the following, we will show that we can achieve a linear lower bound for both real and complex field.

4.1 The real case

The following result is a consequence of Proposition 3.1

Proposition 4.2. *Consider any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$, with $a_i \in \mathbb{R}$ satisfying the Pólya condition. Then we have $\dim\langle F \rangle \geq \lfloor \frac{s+4}{3} \rfloor$.*

Proof. Assume that $e_1 \geq e_2 \geq \dots \geq e_s$ and consider the family $G := \{(x - a_i)^{e_i} : 1 \leq i \leq t\} \subset F$ with $t := \lfloor \frac{s+4}{3} \rfloor$. Since F satisfies the Pólya condition, we have that $e_{s-i} \geq i$ and, hence, $e_i \geq s - t$ for all $i \in \{1, \dots, t\}$. The inequality $s - t \geq 2t - 4$ holds, thus we conclude by Proposition 3.1 that the elements of G are linearly independent and $\dim\langle F \rangle \geq \dim\langle G \rangle = |G| = t$. \square

With more work, we can achieve a better lower bound.

Consider a family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$. By a *sequence in the node* $a \in \mathbb{R}$ we mean a maximal interval of integers O such that for all $e \in O$, the element $(x - a)^e$ belongs to F . A sequence O with an odd size is naturally called an *odd sequence*. The following result is just a restatement of [6, Corollary 9.(ii)]; this result was obtained by transforming the problem of linear independence of shifted powers into an equivalent problem in Birkhoff interpolation, and then applying a celebrated result of Atkinson and Sharma [2] concerning real Birkhoff interpolation (see [15, Theorem 1.5]).

Corollary 4.3. *Consider a family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$, with $a_i \in \mathbb{R}$ satisfying the Pólya condition and set $d := \max(e_i)$. If every odd sequence O in any node a_i satisfies that $\max(O) = d$, then the elements of F are linearly independent.*

Proposition 4.4. *Consider any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$, with $a_i \in \mathbb{R}$ satisfying the Pólya condition. Then we have $\dim\langle F \rangle \geq \lfloor \frac{s}{2} \rfloor + 1$.*

Proof. Let us denote $d = \max(e_i)$. If there is only one i such that $d = e_i$ we have $\dim\langle F \rangle = \dim\langle F \setminus \{(x - a_i)^{e_i}\} \rangle + 1$. So, without loss of generality we assume that there are at least two e_i 's equal to d . Let O_1, \dots, O_k be all the odd sequences such that $d \notin O_i$; notice that $k \leq s - 2$. We denote $m_i = \min(O_i)$, $M_i = \max(O_i)$ and call b_i the corresponding node; we order O_1, \dots, O_k so that $m_1 \leq \dots \leq m_k$. We claim that $F' := F \setminus \{(x - b_i)^{m_i} : 1 \leq i \leq \lceil k/2 \rceil\}$ yields a linearly independent family. Since the size of F' is $s - \lceil k/2 \rceil \geq s - \lceil (s - 2)/2 \rceil = \lfloor \frac{s}{2} \rfloor + 1$, this will prove that $\dim\langle F \rangle \geq \dim\langle F' \rangle \geq \lfloor \frac{s}{2} \rfloor + 1$.

More precisely, we are going to prove that the family $F' \cup \{(x - b_i)^{M_i+1} : \lceil k/2 \rceil + 1 \leq i \leq k\}$ is linearly independent. Indeed, by construction of F' , every odd sequence O in this family satisfies $\max(O) = d$; this is because we have removed an element from $O_1, \dots, O_{\lceil k/2 \rceil}$ and added one to $O_{\lceil k/2 \rceil + 1}, \dots, O_k$, which converts all these odd sequences into even sequences. Moreover, the new set also satisfies the Pólya condition since for every $1 \leq i < j \leq k$, we have $m_i \leq m_j \leq M_j + 1$, so removing a shifted power of exponent m_i and adding one of exponent $M_j + 1$ can never cause a violation of the Pólya condition. By Corollary 4.3 we are done. \square

It is worth pointing out that this result does not only bound the dimension of $\langle F \rangle$ but also shows how to explicitly obtain a linearly independent subset of F of size at least $\lfloor \frac{s}{2} \rfloor + 1$. This will not be the case in the complex setting, where we will obtain a lower bound on the dimension of $\langle F \rangle$ but we will not provide an explicit linearly independent subset of F of size $\Omega(s)$.

We do not know if the above bound is sharp. In the following example we exhibit a family F of s shifted powers that satisfy the Pólya condition and we prove that $\dim\langle F \rangle = (3s - 1)/4$.

Lemma 4.5. *Let $d \in \mathbb{Z}^+$ such that $d \equiv 2 \pmod{4}$ and consider $F_1 := \{x^i : i \text{ odd and } i < d\}$ and $F_2 := \{(x + 1)^i, (x - 1)^i : i \text{ even and } (d + 2)/2 \leq i \leq d\}$. The family $F := F_1 \cup F_2$ has $d + 1$ elements, satisfies the Pólya condition and $\dim\langle F \rangle = (3d + 2)/4$.*

Proof. It is easy to see that $|F_1| = d/2$, $|F_2| = (d + 2)/2$ and that F satisfies the Pólya condition. We remark that F has an element of degree i for all $i \in \{1, 3, \dots, d/2, (d/2) + 1, \dots, d\}$. This implies $\dim\langle F \rangle \geq (3d + 2)/4$. Moreover, for every i even such that $(d + 2)/2 \leq i \leq d$ we observe that $(x + 1)^i - (x - 1)^i = \sum_{\substack{j < i \\ j \text{ odd}}} 2 \binom{i}{j} x^j \in \langle F_1 \rangle$. Hence, we can combine every pair elements of F_2 of the same degree so that the combination can be expressed as a linear combination of the elements of F_1 . This proves that $\dim\langle F \rangle \leq |F| - \frac{|F_2|}{2} = d + 1 - \frac{d+2}{4} = \frac{3d+2}{4}$.

4.2 The complex case

In the complex case, our results rely on Corollary 2.5 and on a good choice of parameters in Proposition 2.2.

Proposition 4.6. *Fix an integer $p \in \mathbb{N}$ and $\alpha > 0$ and consider any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq p\}$ with $e_i \geq \alpha p$ for all i . Then we have $\dim \langle F \rangle > (1 + \alpha - \sqrt{\alpha^2 + 1})p$.*

Proof. We extract a basis $G \subseteq F$ of $\langle F \rangle$. We set $s := |G|$, $\gamma := (1 + \alpha - \sqrt{\alpha^2 + 1})p$ and we assume for contradiction that $s \leq \gamma$. We take $t = s$, $k = \alpha p$ and $l = p - 1$, we claim that the inequality in Proposition 2.2 is satisfied. Indeed, according to Proposition 2.2, it suffices to prove that the polynomial function

$$\begin{aligned} p(S) &= S^2 - (2l + 2k + 3)S + 2(k + 1)(l + 1) \\ &= S^2 - (2p + 2\alpha p + 1)S + 2p(\alpha p + 1) \end{aligned}$$

is positive for $S = s$. We will prove that $p(S) > 0$ for all $S \in [0, \gamma]$. For this purpose, we consider

$$q(S) := p(S) + S - 2p = S^2 - (2p + 2\alpha p)S + 2\alpha p^2;$$

it is straightforward to check that $q(S) \geq 0$ for all $S \in [0, \gamma]$ since γ is the smallest root of q . We conclude that $p(S) > 0$ in $[0, \gamma]$ because the following inequalities hold for every $S \in [0, \gamma]$:

$$p(S) = q(S) - S + 2p \geq q(S) - \gamma + 2p = q(S) + (\sqrt{\alpha^2 + 1} - \alpha + 1)p > q(S).$$

Hence, by Proposition 2.2, there exists an SDE($s, k', p - 1$) with $s \leq k' \leq \alpha p$ satisfied by all the elements of G . Since G is a basis of $\langle F \rangle$, it follows that not only the elements of G are solutions of this SDE, but also the elements of F . Since for any $i \in [1; p]$ we have $e_i \geq \alpha p \geq k'$, it follows from Corollary 2.5 that $p \leq \deg P_{k'} \leq p - 1$, a contradiction. \square

Corollary 4.7. *Fix an integer $s \in \mathbb{N}$ and consider any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ with $e_i \geq s$ for all i . Then we have $\dim \langle F \rangle > (2 - \sqrt{2})s$.*

We now extend Proposition 4.6 by allowing small exponents.

Theorem 4.8. *For any family $F = \{(x - a_i)^{e_i} : 1 \leq i \leq s\}$ satisfying the Pólya condition, we have $\dim \langle F \rangle \geq (1 - \frac{\sqrt{2}}{2})(s - 1)$.*

Proof. We start by dropping the $\lceil \frac{s}{2} \rceil$ terms with least exponents and call F' the resulting family. By construction we have $|F'| = s - \lceil \frac{s}{2} \rceil = \lfloor \frac{s}{2} \rfloor$, and because F satisfies the Pólya condition, every element $(x - a)^e \in F'$ must verify $e \geq \lceil \frac{s}{2} \rceil$. Thus the family F' satisfies the hypothesis of Corollary 4.7, implying that $\dim \langle F' \rangle \geq (2 - \sqrt{2}) \lfloor \frac{s}{2} \rfloor$. The result follows since $\langle F' \rangle$ is a subset of $\langle F \rangle$. \square

5 Conclusion

We have proposed two conjectures concerning linear independence of shifted powers; we have proved their counterparts for the field of real numbers and have given some steps towards potential proofs over \mathbb{C} . The proof of Conjecture 1.7, apart from being interesting by itself, has nice consequences in arithmetic complexity: it implies a linear lower bound on the number of shifted powers needed to represent a polynomial of degree d . We believe that both of them are true; however, we have the feeling that a tool different from shifted differential equations should be used to prove them.

We have also provided bounds on the dimension of the vector space spanned by a family F of shifted powers that satisfy the Pólya condition. The lower bounds for the field of complex numbers given in Section 4.2 are nonconstructive in the sense that they do not pinpoint a linearly independent subset of F of cardinality equal to our lower bound on $\dim F$ (but they of course imply the existence of such a subset). It would be interesting to obtain a constructive proof. This may be related to the problem of obtaining a “good” sufficient condition for the linear independence of shifted powers over \mathbb{C} .

To our knowledge, the family F of real or complex polynomials that satisfies the Pólya condition and that spans a vector space with the least dimension is the one we provide in Lemma 4.5. It would be interesting to improve the bounds we provide or to show that they are tight.

Acknowledgements

The authors are supported by ANR project CompA (code ANR-13-BS02-0001-01). I. García-Marco was also partially supported by MTM2016-78881-P. We thank the anonymous referees for suggesting several improvements in the presentation of this paper.

References

- [1] V. Arnold. *Lectures on Partial Differential Equations*. Springer, 2004.
- [2] K. Atkinson and A. Sharma. A partial characterization of poised Hermite-Birkhoff interpolation problems. *SIAM Journal on Numerical Analysis*, 6(2):230–235, 1969.
- [3] M. Bôcher. The theory of linear dependence. *The Annals of Mathematics*, 2(1/4):81–96, 1900–1901.
- [4] A. Bostan and P. Dumas. Wronskians and linear independence. *The American Mathematical Monthly*, 117(8):722–727, 2010.
- [5] G. Comas and M. Seiguer. On the rank of a binary form. *Foundations of Computational Mathematics*, 11(1):65–78, 2011.
- [6] I. García-Marco and P. Koiran. Lower bounds by Birkhoff interpolation. *J. Complexity*, 39:38–50, 2017.
- [7] I. García-Marco, P. Koiran, and T. Pecatte. Reconstruction algorithms for sums of affine powers. *arXiv preprint arXiv:1607.05420 [cs.CC]*, 2016. Conference version: *Proc. ISSAC 2017*, pp. 317–324, <http://dx.doi.org/10.1145/3087604.3087605>
- [8] J. H. Grace and A. Young. *The algebra of invariants*. Cambridge Library Collection. Cambridge University Press, Cambridge, 2010. Reprint of the 1903 original.
- [9] A. Iarrobino and V. Kanev. *Power Sums, Gorenstein Algebras, and Determinantal Loci*. Springer, 1999.
- [10] N. Kayal, P. Koiran, T. Pecatte, and C. Saha. Lower bounds for sums of powers of low degree univariates. In *Proc. 42nd International Colloquium on Automata, Languages and Programming (ICALP 2015), part I*, LNCS 9134, pages 810–821. Springer, 2015. Available from <http://perso.ens-lyon.fr/pascal.koiran>.
- [11] D. E. Knuth. *The Art of Computer Programming, Volume 4, Fascicle 4: Generating All Trees-History of Combinatorial Generation (Art of Computer Programming)*. Addison-Wesley Professional, 2006.
- [12] M. Krusemeyer. Why does the Wronskian work? *The American Mathematical Monthly*, 95(1):46–49, 1988.

- [13] Y. N. Lakshman and B. D. Saunders. Sparse shifts for univariate polynomials. *Applicable Algebra in Engineering, Communication and Computing*, 7(5):351-364, 1996.
- [14] G. Lorentz and S. Riemenschneider. Probabilistic approach to Schoenberg's problem in Birkhoff interpolation. *Acta Mathematica Hungarica*, 33(1-2):127-135, 1979.
- [15] G. G. Lorentz, K. Jetter, and S. D. Riemenschneider. *Birkhoff interpolation*, volume 19 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1984.
- [16] R. Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Version 3.1.3, 2017. github.com/dasarpmar/lowerbounds-survey/releases.
- [17] M. Yoshida, Fuchsian differential equations. With special emphasis on the Gauss-Schwarz theory. Aspects of Mathematics, E11. Friedr. Vieweg & Sohn, Braunschweig, 1987.